REVIEW OF ENVIRONMENTAL MONITORING METHODS: SURVEY DESIGNS

TRENT L. MCDONALD
Western EcoSystems Technology, Inc., Cheyenne WY, 82001, e-mail: tmcdonald@west-inc.com

(Received 6 November 2001; accepted 20 August 2002)

Abstract. During the past decade and a half, environmental monitoring programs have increased in number and importance. Large scale environmental monitoring programs often present design difficulties because they tend to measure many (sometimes hundreds) of parameters through space and time. This paper reviewed and summarized one important component of environmental monitoring programs, the statistical survey design. Survey designs used for long-term monitoring programs lasting multiple ($\geq$ 3) occasions were reviewed, paying special attention to those published after 1985. During this review, two key components of the overall survey design were identified. The first key component was the membership design. Groups of population units sampled the same occasion were called panels here, and the membership design specified which units were members of which panels. The second component was the revisit design that specified when panels were to be revisited. Membership designs varied, but some form of simple random or systematic design was popular. Among revisit designs, four basic patterns were found in the literature and their relative strengths and weaknesses were summarized. To efficiently discuss revisit designs, a new unified short-hand notation was proposed and adopted.

Keywords: finite population, change, literature summary, review, sample

1. Introduction

When designing a large scale statistical survey, researchers in all fields must grapple with a number of difficult issues. These issues include the definition of the sample unit, how to disperse the sample through space and time, how to construct an accurate sampling frame, whether to conduct equiprobable sampling or variable probability sampling, whether to stratify and how, how to account for frame errors, and how to account for non-response and measurement errors. They must anticipate the presence and effects of spatial correlation, anticipate yearly (temporal) variation, whether to estimate status or trends or both, etc. In making decisions about these issues, compromise is often called for because researchers must balance their natural desire to revisit the same sample units to improve estimation of trend with their desire to visit unmeasured sample units to improve estimation of status.

Taken together, these and other survey design issues can easily overwhelm and confuse a survey designer, even experienced statisticians. To add to the confusion, there has been an abundance of statistical research in the area of environmental survey design and trend detection in the last decade. Non-standard terminology and

Environmental Monitoring and Assessment **00:** 1–16, 2002.

varying notation in this rapidly growing topic have added to researcher confusion. It was clear that a non-technical review and summary of the statistical literature was needed.

In this paper, one important component of the overall survey design was summarized in a non-technical manner. Specifically, general notation, useful revisit schemes, and useful sample designs to populate survey panels published in the recent literature on ecological trend methods were reviewed and summarized. Certain other aspects of survey design, such as sample unit definition, frame construction, analyses, etc, are important topics but were beyond the scope of this review (see Kish (1965), Cochran (1977), Krishnaiah and Rao (1988), and Sarndal *et al.* (1992)). This article is intended as a general guide to environmental survey design for biologists, ecologists, and managers who are faced with the daunting task of designing a large scale environmental survey. Details necessary to actually carry out such a survey will need to be gleaned from the references contained herein.

It is worthwhile noting a few essential references for environmental monitoring methods that were found during the course of this review. The first two are the proceedings of two conferences on large scale survey methods over time. The first conference, entitled Ecological Resource Monitoring: Change and Trend Detection, was held in 1996 and selected papers appear in a special issue of Ecological Applications (1998, Vol 8, No. 2) (Edwards, 1998, contains a summary). The second conference, entitled Environmental Monitoring Surveys Over Time, was held in 1998 and selected papers were published in a special issue of the Journal of Agricultural, Biological, and Environmental Statistics (1999, Vol 4, No. 4) (Olsen and Smith, 1999, contains a summary). In addition, a statistician may find Krishnaiah and Rao (1988) useful. Example studies can also provide valuable information on study design. Examples of current large scale environmental surveys include the National Resources Inventory (NRI) (Nusser and Goebel, 1997; Nusser *et al.*, 1998), the Forest Inventory and Analysis (FIA) (McRoberts and Hansen, 1999; Reams and Deusen, 1999; Moisen and Edwards, 1999; Leatherberry *et al.*, 1995), the Forest Health Monitoring survey (FHM) (Eager *et al.*, 1991), the National Wetlands Inventory (NWI) (Ernst *et al.*, 1995), and, in the recent past, the Environmental Monitoring and Assessment Program (Overton *et al.*, 1990).

In Section 2, study objectives and several definitions of trend are given in order to identify advantageous survey designs for each objective. Section 3 proposes unified notation for the rotation of sample effort through time among groups of sample units called panels. The rotation of sample effort among survey panels will be called the *revisit* design, and four basic revisit designs will be discussed in Section 4. The method used to populate survey panels will be called the *membership design*, and will be treated in Section 5.

2. Study Objectives and Definition of Trend

Most studies over extended time periods are designed to estimate status (mean levels) and detect trends in environmental parameters. However, it became obvious during this review that the definition of trend varied according to study objectives. This section gives some definitions of trend in order to clarify what is being estimated and, consequently, study objectives. It will also be useful to point out which survey designs are best suited to satisfy certain objectives.

Duncan and Kalton (1987) list the following objectives for a survey conducted over time: (1) to estimate population parameters at distinct time points, (2) to estimate population parameters averaged across a period of time, (3) to measure net change, (4) to measure components of individual change, (5) to aggregate data for individuals over time, (6) to measure frequency, timing, and duration of events, and (7) to accumulate samples over time, especially rare samples. Of these objectives, (3) and (4) relate to the estimation and detection of trend.

In objective (4), measurement of *individual change* means measurement of the change experienced by an individual or particular member of the population. Individual change can be further categorized as one of three types. The first type is *gross change* and is defined to be the change in the response of a particular population unit. For example, gross change happens when the pH value of a particular lake increases or decreases. Gross change also happens if all rivers in a collection of rivers have higher levels of suspended sediment. The second type of individual change is *average gross change*. Average gross change is only defined for units in the population for two or more time periods (e.g., additions to the population are excluded). For example, average gross change might be the average difference in plant density on a set of quadrats in a study area. The third type of individual change is *instability* and is defined as the variance of responses from individual population units. For example, instability might be estimated as the variance of waterfowl numbers on a particular wetland.

In objective (3), measurement of *net change* means measurement of total change in a parameter arising from all sources. Sources that might cause net changes in a parameter include immigration or emigration of sample units and any component of individual change. Net change can be thought of as change in the mean or total response. It should also be clear that individual change can happen without causing net change because, for example, responses from some units can go up while responses from others go down. Suppose that the fish in a population of stream segments shift their spatial distribution during some time period, but none die or emigrate out of the population. In this case, individual stream segments are experiencing trend (i.e., gross trend), but overall the number of fish remains constant so that the population experiences no net trend. If a population does not add or delete units between two time periods, it can only experience net change as a result of individual change. If responses are constant between two time periods, the population can only experience net change through the addition or deletion of population units.

3. Unified Notation for Revisit Designs

It became apparent during this review that lack of a unified notation, varying descriptions, and varying names for the same schedule of revisits to a population unit were all barriers to a clear understanding of the overall survey design. Duncan and Kalton (1987, p. 99) acknowledged this when they stated, “Unfortunately there is no widely-accepted terminology for these designs...”. This section proposes and adopts a simple and flexible short-hand notation that will facilitate better understanding of revisit patterns. First, however, some definitions are needed.

In what follows, the term *panel* will be defined to be a group of population units that are always all sampled during the same sampling occasion or time period. A population unit is allowed to be a member of two different panels under this definition. The pattern of visits to a panel will be called the *revisit design*, or more generically, the *revisit design* will refer to the plan by which population units are visited and sampled through time. The way in which units of the population become members of a panel will be called the *membership design*. For example, a membership design might dictate that panel members be chosen using independent simple random sampling, while the revisit design might dictate that each panel be visited only once. Under this membership design (independent simple random sampling), an alternative revisit design might dictate that the first sample be visited every occasion, but all other samples be visited only once. In fact, the revisit design has the option of specifying any advantageous revisit cycle for members of each panel, regardless of how those members were chosen. For this reason, the revisit design is independent and distinct from the membership design.

The unified notation proposed here for revisit designs extends the “4-8-4” notation used by Binder and Hidiroglou (1988) and others. In the proposed notation, a panel revisit scheme will be represented by a string of digits, separated by dashes, and enclosed in brackets. The number and value of the digits will be assigned according to the following rules: (1) every digit in an odd numbered location (i.e., the 1st, 3rd, 5th, etc.) will be the number of consecutive occasions that a panel is sampled before it is rotated out of the sample and not visited for a period of time (0, meaning “not visited”, is allowed), (2) every digit in an even numbered location (i.e., the 2nd, 4th, 6th, etc.) will be the number of consecutive occasions that a panel is not sampled before it is rotated back into the sample and visited again for some period of time (0, meaning “not rotated out of the sample", is allowed), (3) there must be an even number of digits in the descriptive string, (4) the letter ’n’ in place of a digit in an even numbered location denotes that the panel is never revisited again, (5) if the last digit in a string is not an ’n’, the revisit scheme defined by the preceding digits is repeated indefinitely, (6) different revisit schemes for different sets of panels are separated by a comma, (7) if the number of panels with the specified revisit design is not the default (see below), the number of panels enduring the specified revisit design can be denoted by a super-script following a string of digits that are enclosed in parentheses, and (8) an unspecified number of occasions can be represented by a letter, other than 'n', in place of any digit.

In this notation, [1-n] implies that a panel should be visited once and never again. The notation [1-0] implies that a panel should be visited once and never leave the sample (i.e., always revisited). The notation [4-8-4-n] implies that a panel should be sampled for four consecutive sample periods, rotated out of the sample for eight periods, rotated back into the sample for four periods, and never revisited after that (this scheme was used in the U.S. Census Bureau's Current Population Survey (Binder and Hidiroglou, 1988)). Rule (6) accommodates so-called *split-panel* designs, so that [1-0,1-n] denotes a simple split panel design (see Fuller (1999) and Duncan and Kalton (1987) regarding this particular design) in which one panel is always revisited while the remaining panels are sampled for one period and then dropped. The [$(1\text{-}5)^3$] revisit design implies a new panel will be visited during each of the first three occasions, no panels will be sampled for three occasions, the three original panels will be visited in order over the next three occasions, sampling will again cease for three occasions, and so on. The [x-y] revisit design specifies a rotating panel design where panels are sampled for x periods, not sampled for the next y periods, sampled again for the next x periods, not sampled for the next y periods, etc. The notation [a-b,c-d] represents a general split panel design with two revisit plans, one plan for each of two sets of panels.

The default number of panels enduring a particular revisit plan can be derived from the notation. If a string of revisit digits does not end with an 'n', the number of panels enduring that revisit plan is equal to the sum of the first odd-even position pair in which the odd member is non-zero. For example, there is one panel implied by the [1-0] design, 12 panels implied by the [4-8-1-11] scheme, and three panels implied by the [0-3-1-2] plan. If a super-script is present, the number of panels is given by the super-script. For example, [$(1\text{-}2)^1$,$(2\text{-}1)^2$] specifies one panel be visited once every three occasions, and two panels be visited twice every three occasions (for a total of 3 panels). If a string of revisit digits ends with a 'n' and does not have a super-script, the revisit scheme is repeated on new panels indefinitely. Consequently, there are an infinite (or very large) number of panels with that particular revisit scheme.

The number of occasions necessary to complete one full cycle can also be derived from the notation. Sampling is said to complete one full cycle when all panels that will ever be sampled are visited an equal number of times. If a revisit design does not end with an 'n', the number of occasions required to visit every panel an equal number of times is the sum of all digits in the string. If a super-script is present, the number of occasions in a complete cycle is the super-script squared. If a design does not have a super-script and ends with 'n', the notion of a sampling "cycle" does not exist, even though it is possible that at some point all panels will

TABLE I

Example startup options for the [3-n] and [3-2] revisit designs. By convention, if a startup plan is not specified for a [x-y] design, an equal number of panels are sampled each occasion (i.e., Options 2 or 3). Layout of this table adapted from Urquhart et al. (1998), Urquhart and Kincaid (1999), Fuller (1999), and others.

Design [3-n]								**Design [3-2]**								
	Sample Occasion								Sample Occasion							
Panel	1	2	3	4	5	6	7	Panel	1	2	3	4	5	6	7	
	Option 1								Option 1							
1	X	X	X					1	X	X	X			X	X	
2		X	X	X				2		X	X	X			X	
3			X	X	X			3			X	X	X			...
4				X	X	X		4				X	X	X		
5					X	X	X	5					X	X	X	
⋮							⋱									
	Option 2								Option 2							
1	X							1	X			X	X	X		
2	X	X						2	X	X			X	X	X	
3	X	X	X					3	X	X	X			X	X	...
4		X	X	X				4		X	X	X			X	
5			X	X	X			5			X	X	X			
⋮					⋱											
									Option 3							
								1	X	X	X			X	X	
								2		X	X	X			X	
								3			X	X	X			...
								4	X			X	X	X		
								5	X	X			X	X	X	

be sampled an equal number of times. For example, the [1-0] design completes a cycle in one occasion, the [3-2] design completes a cycle in 5 occasions, the [4-8-1-11] design completes a cycle in 24 occasions, and the [(2-1)^2] design completes a full cycle in 4 occasions.

Several startup options exist for all revisit designs, and the notation proposed in this section does not specify which option is used. For revisit designs with two digits (i.e., [x-y]), one startup option specifies that all panels be sampled the same number of occasions (Option 1, Table I). Another startup option for [x-y] designs specifies that the same number of panels be sampled each period (Option 2 or 3, Table I). After the first few periods, the same number of panels will be sampled each period under all options. Unless funding is incremental or the first few occasions are viewed as "pilot" occasions, it is usually best to sample the same number of panels each occasion because required funding levels are constant. By convention, if the startup option is not specified for a [x-y] design, it will be understood that the number of panels sampled each time period is constant (i.e., Option 2 or 3 will be assumed).

For revisit designs with four digits (i.e., [a-b-c-d]) and revisit designs with a super-script, it is generally not possible to sample the same number of panels on all occasions. Convention for these designs will be that all panels should be sampled an equal number of times (Option 1).

Urquhart and Kincaid (1999) used the terms "always revisit", "never revisit", "rotating panel", "augmented serially alternating", and "partially augmented serially alternating" to describe the revisit designs [1-0], [1-n], [5-n], [1-0,1-3], and [1-3,2-2-1-3-1-3,1-3-2-2-1-3,1-3-1-3-2-2] respectively. Fuller (1999) used the terms "independent samples", "pure panel", "rotating panel", and "supplemented panel" to describe designs [1-n], [1-0], [x-n], and [1-0,1-n] respectively. Duncan and Kalton (1987) used the terms "repeated survey", "panel survey", and "rotating panel survey" to describe the [1-n], [1-0], and [x-n] respectively. Kish (1983), Kish (1986), and Duncan and Kalton (1987) use the term "split panel" to describe designs of the form [1-0,x-y]. Skalski (1990) used the term "rotating panel with augmentation" to describe the [1-0,4-n] revisit design.

4. Basic Revisit Designs

There are an infinite number of revisit designs available to sample most populations. This section contains a description of four basic revisit designs found in the literature and their relative strengths and weaknesses.

Using revisit designs to rotate sample effort among panels was proposed as early as the middle of the 20th century. Jessen (1942) studied rotating panel designs [x-y] where simple random sampling determined members of each panel. Design-based estimation, primarily under simple revisit designs and simple random sampling, was further considered by Yates (1949), Patterson (1950), Hansen *et al.* (1953), Eckler (1955), Hansen *et al.* (1955), Rao and Graham (1964), Gurney and Daly (1965), Prabhu-Ajgaonkar (1967), Singh (1968), Agarwal and Tikkiwal (1975), Cochran (1977), Smith (1978), Tikkiwal (1979), and Wolter (1979). Revisit bias, arising from the fact that repeated observations on the same sample unit do not necessarily have the same expected value, was considered by Hansen *et al.* (1955), Gurney and Daly (1965), Bailar (1975), Bailar (1979), Huang and Ernst (1981), and Kumar and Lee (1983). Design-based analyses that estimate trend using relatively simple revisit designs was summarized in the late 1980's by Binder and Hidiroglou (1988) and Duncan and Kalton (1987).

Among all these articles and ones since the late 1980's, four basic revisit schemes were common. These four basic revisit schemes were [1-0], [1-n], [x-y], and [a-b,c-d]. Specific examples of these four designs and two others are portrayed in Table II.

Under design [1-0], the same panel is sampled each occasion. An advantage of this revisit scheme is that planning and survey design work are minimized because measurements are taken from all units in a single panel every occasion. Under this scheme, navigation to and identification of study units after occasion one is usually trivial because units have already been visited. Because every unit is visited every occasion, the [1-0] design is well suited to estimate gross change and other components of individual change (Duncan and Kalton, 1987). In addition, past information from a unit can be used in a number of ways to both strengthen current estimates and detect trend (Binder and Hidiroglou, 1988). This design admits both classical and time series analysis described by Binder and Hidiroglou (1988). Urquhart and Kincaid (1999) found this design to be the most powerful for detecting linear trend among the five they investigated.

A primary drawback of the [1-0] design is the fact that the scheme places a high response burden on members of the chosen panel. High response burden can cause sampled units to "wear out" and either quit responding (i.e., become non-responses) or yield responses that have changed simply because they have been measured several times. Non-responses in particular are problematic and can occur when a particular unit can no longer be located (i.e., failure to keep track of the unit), or because the unit left the population, or because permission was refused. The fact that responses from repeatedly sampled units depend, to some extent, on previous responses and previous collection attempts has been called *conditioning* by Duncan and Kalton (1987) and Fuller (1999). Fuller (1999, p.339) states "...if we go repeatedly to a particular forest or range site, we will change the behavior of the plants and animals on that plot. The conditioning effect of repeated measures on the same unit should not be underestimated." Duncan and Kalton (1987, Section 4) describe methods to account for non-response by re-weighting estimates or imputation.

In conjunction with high response burden, the [1-0] design cannot automatically account for changes in population composition over time. For example, if 100 units were added to the population after the original panel was selected, responses from the panel will not represent these 100 new units. In large scale surveys over extended time periods, the size or composition (or both) of the population is likely to change and study designers usually choose to incorporate a facility that adds new units to the sampled panel at certain times in the future. Comparison of responses from new units to those from the original panel allows investigation (and a potential adjustment) of the effects of changing population composition. When new units are

TABLE II
Schematic and notational representation of seven typical revisit designs.

Panel	Sample Occasion 1	2	3	4	5	6	7	8	9	10
A: Design [1-0]										
1	X	X	X	X	X	X	X	X	X	X
B: Design [1-n]										
1	X									
2		X								
3			X							
4				X						
5					X					
6						X				
7							X			
8								X		
9									X	
10										X
⋮										
C: Design [2-n]										
1	X									
2	X	X								
3		X	X							
4			X	X						
5				X	X					
6					X	X				
7						X	X			
8							X	X		
9								X	X	
10									X	X
⋮										
D: Design [2-3]										
1	X	X				X	X			
2		X	X				X	X		
3			X	X				X	X	
4				X	X				X	X
5	X				X	X				X

Panel	Sample Occasion 1	2	3	4	5	6	7	8	9	10
E: Design [1-0,1-n]										
1	X	X	X	X	X	X	X	X	X	X
2	X									
3		X								
4			X							
5				X						
6					X					
7						X				
8							X			
9								X		
10									X	
11										X
⋮										
F: Design [1-0,2-2-2-n]										
1	X	X	X	X	X	X	X	X	X	X
2	X	X			X	X				
3		X	X			X	X			
4			X	X			X	X		
5				X	X			X	X	
6					X	X			X	X
7						X	X			X
8							X	X		
9								X	X	
10									X	X
11										X
⋮										
G: Design [(1-1-1-n)^2,(0-4-1-1-1-n)^2,...]										
1	X		X							
2		X		X						
3					X		X			
4						X		X		
5									X	
⋮										

added, the revisit design is not pure [1-0], but is a split panel design of the form [1-0,x-y] where the [1-0] component experiences (perhaps random) losses.

The second basic revisit design scheme, [1-n], dictates that each panel be visited only once. The advantages of this design include extremely low response burden, quick accumulation of information from a large number of population units, and the ability to automatically take account of changes in population composition. Even though sampled units will be changed in the measurement process, the conditioning effect mentioned above cannot occur because units are never revisited. Individual units can be revisited under this design if the sample design allows the same unit to be a member of two or more panels. Fuller (1999) states that the [1-n] design is a good procedure for locating rare items. Both Fuller (1999) and Duncan and Kalton (1987) note that [1-n] can be used to estimate net change. Duncan and Kalton (1987) state that [1-n] can be used to meet design objectives (1), (2), (3), (6), and (7) above. Urquhart and Kincaid (1999) found that [1-n] had higher power to detect trend than a [5-n] design with (1/5)-th of the sample allocated to each panel. Urquhart *et al*. (1998) note that for estimation of population status, the [1-n] design is eventually the most efficient design because it incorporates the most number of unique sites; however, they also note that augmented or split-panel designs of the form [1-0,1-x] or [1-x,a-b-c-d-e-f] are also efficient for estimating status over intermediate time periods (5-15 occasions).

The primary disadvantage of the [1-n] design is that it does not allow estimation of gross, or individual, change (Fuller, 1999; Duncan and Kalton, 1987). Furthermore, [1-n] generally has lower statistical efficiency for estimation of net change than do other designs (Fuller, 1999). The [1-n] is also not connected in the experimental design sense, complicating application of standard linear models (Urquhart *et al*., 1998; Urquhart and Kincaid, 1999). Duncan and Kalton (1987) state that the [1-n] design does not yield data sufficient to satisfy objectives (4) and (5), and for objective (6) the design is subject to telescoping errors. Urquhart *et al*. (1998) and Urquhart and Kincaid (1999) found that the [1-n] and [5-n] designs generally had low power to detect trends among those designs they investigated.

In general, the other two designs, [x-y] and [a-b,c-d], are intermediate between the [1-0] and [1-n] designs and generally offer compromises among the advantages and disadvantages of each. Another revisit design that is intermediate between [x-y] and [a-b,c-d] is the [x-y-z-n] design. The [x-y-z-n] design offer both extended time between measurements relative to [x-y] and the same rate of accumulation of new units. Both the [x-y] and [x-y-z-n] designs are good for estimation of both status and trend (Binder and Hidiroglou, 1988). Split panels of the form [a-b,c-d] offer similar compromises to [x-y] and [x-y-z-n] designs, but can be designed so that the plan is connected in the experimental design sense and therefore can be analyzed using standard linear model theory (Urquhart and Kincaid, 1999). If the split panel design has a component that is always revisited (i.e., [1-0,x-y]), the response burden on certain units will be high, but the response burden on other units will be low. Urquhart *et al*. (1998), Urquhart and Kincaid (1999), and Breidt and

Fuller (1999) found that split panels generally had high power for both estimation of trend and status.

5. Membership Designs

It was common for membership designs to be either not mentioned or described inadequately in the articles reviewed here. Without specific descriptions, it was left for the reader to infer that population units might have been collected without a specific probability sample, perhaps using judgment or haphazard samples. Judgment samples contain sample units that were hand-picked by researchers, usually to be "representative". Haphazard samples are samples that have been collected without a defined protocol or directed effort. Both types of samples are non-probability samples in the sense that randomization is completely absent from the process by which population units are picked. While there can be a place for judgment and haphazard sampling in science, there is a limit to what can be reasonably inferred from haphazard or judgment samples (Olsen *et al.*, 1999), and both have been a source of considerable controversy in the past and their use has lead to some famous miscalculations (Edwards, 1998).

The problem with judgment or haphazard samples is that inference to the larger population requires assumptions about the behavior of units and the responses of non-sampled units. A common assumption is that responses on sampled units are "representative" of responses on non-sampled units. However, "representative" is difficult to define and even if a reasonable definition can be found, "representativeness" can not be verified in the vast majority of studies. Moreover, there is almost surely some bias introduced with each additional assumption, and usually the severity of this bias cannot be estimated (Olsen *et al.*, 1999). To be useful and accurate, all assumptions should be validated and make ecological sense (Olsen *et al.*, 1999).

A common probability based membership design in the literature was simple random sampling without replacement. Using simple random sampling, all possible samples were equally likely to occur with probability $1/\binom{N}{n} = n!(N-n)!/N!$, where N was population size and n was sample size. For panel membership, it is possible to either take independent simple random samples for each panel, or take a single simple random sample and divide it among panels. For example, if k panels containing n units each are to be sampled, it is possible to draw k independent samples for the members of each panel, or draw one simple random sample of size kn, assigning the first n to panel 1, the second n to panel 2, and so on. Dividing a single sample among panels results in dependent samples in each panels because units in one panel are known not to be members of another panel (assuming without replacement sampling).

Another common probability based membership design was systematic sampling. In environmental monitoring studies where population units were often plots or points on a landscape, systematic samples were generally some type of 2-dimensional geographic grid sample, although samples could also be taken systematically over other variables (e.g., elevation, distance to roads, etc.). Like simple random sampling, independent systematic samples can be taken for the members of each panel, or a large systematic sample can be taken and divided among the panels. Dividing a single systematic sample among panels gives rise to so-called *interpenetrating* sample designs (Reams and Deusen, 1999). If k panels containing n units each are to be constructed, an interpenetrating sample designs specifies first that a systematic sample of size kn be taken from the population. Then, members of panel 1 are units $1, k+1, 2k+1, \ldots, (n-1)k+1$ of the sample. Members of panel 2 are units $2, k+2, 2k+2, \ldots, (n-1)k+2$ of the sample, and so on. These types of membership designs are interpenetrating in the variable(s) used to order the systematic sampling, usually geographic space for environmental monitoring studies.

The prime advantage of both simple random sampling and systematic sampling is that statistical inferences to the entire population are straight forward; however, this advantage is only apparent after data have been collected and analysis begins. Both designs include units with constant probability. Technically, this means that every unit in the population is included in the sample with probability n/N under both designs. Simple random sampling also has the characteristic that the presence of a unit in the sample does not preclude other units from being in the sample (i.e., every pair of units is included with probability $n(n-1)/N(N-1)$). With this characteristic, simple random sampling nearly satisfies the *i.i.d.* (independent and identically distributed) assumption required by most classical statistical analyses. As a result, standard analyses such as t-tests and regression can be applied without deleterious effects. Systematic sampling, which does not allow some pairs of units to be sampled, nonetheless satisfies the *i.i.d.* to an extent that classical statistical analyses can usually be applied without adverse effects on inferences. If a membership design does not result in all units being included with equal probability, analyses would be required to account for the variable inclusion probabilities through use of sample weights. Many times, analyses that account for unequal probabilities of inclusion are difficult and should only be carried out by a qualified statistician.

The main disadvantage of both simple random sampling and systematic sampling is the expense involved in locating, traveling to, and collecting data on dispersed and often far-flung units. The popularity of simple random sampling also suffers because a poor randomization can result in units being sampled in a clump, or in sample units that are close to one another based on some measure (like distance). Even though the probability of obtaining a poor randomization is often remote, many researchers are unwilling to take this chance. In some situations, another disadvantage is that simple random and systematic membership designs sample population groups in proportion to their prevalence. Simple random and systematic sampling therefore will probably result in panels without members of rare groups unless sample size is large. If members of rare population groups are highly desired in some panels, alternatives to pure simple random and systematic membership designs should be considered. In these cases, a split-panel design wherein one split specifically targets members of the rare groups might be considered.

6. Discussion

Unfortunately, no single survey design was found to be optimum for all potential objectives of a large scale environmental monitoring project. The choice of survey design depends heavily on the importance placed on specific objectives and size of the study's budget, and consequently a uniform recommendation for all studies cannot be made. However, consensus opinion among the reviewed articles appeared to be that some sort of split panel design had the best chance of satisfying the sometimes competing objectives inherent in many environmental monitoring projects. Split panel designs can be designed to produce relatively rapid accumulation of new population units, as well as adequate numbers of revisits to previously sampled units. Among the alternatives, [1-n] accumulated new units faster than other designs and was good for estimation of status, but [1-0] and [a-b,c-d] were found to be more efficient for estimation of trend. The [1-0] design was found to be a good design for estimating gross trend (i.e., components of individual change), but [1-0] has high response burden, high potential for responses to change simply because measurements were taken multiple times, and lacked an automatic facility to account for changes in population composition. Beyond recommending split panel designs, further recommendations regarding specific revisit cycles cannot be made. The best revisit cycle for a given component of a split panel design will need to be derived for each application after consideration of the biology or life history of the organism(s) being monitored.

Relatively little attention was given in the literature to the question of adequate sample sizes and allocation of sample effort between components of a split panel revisit design. The power analyses of Urquhart and Kincaid (1999) and Urquhart *et al.* (1998) were exceptions, as well as Binder and Hidiroglou (1988), Breidt and Fuller (1999), Fuller (1999) who discussed allocation of sample effort between components of the split panel design [1-0,1-n]. The latter three articles generally concluded that more than 50% of the total sample should be allocated to the [1-0] part of a split panel design if trend is the primary interest. If both trend and status are of interest, allocating exactly 50% of the sample to the [1-0] (revisit) panel and 50% to the [1-n] (rotating) panel is a reasonable compromise (Breidt and Fuller, 1999). Optimum or advantageous allocation of effort among panels of a split panel design is unknown and probably much different than 50-50 under other revisit plans. More research is needed on optimum or advantageous allocation of sample effort between the panels of a split-panel design. For example, how much sample effort should be placed in each part of a [1-1,1-2] or [1-0,1-2,1-5-1-n] design?

When it was possible to determine membership design, simple random sampling and systematic sampling were popular designs to populate panels. With systematic designs, some studies advocated drawing a large sample and allocating it among panels to arrive at an interpenetrating sample.

Use of the short-hand notation of this article, acknowledgement that revisit and membership designs are separate entities, and identification of common revisit and membership designs will foster clear communication of overall survey designs and facilitate expanded statistical research into the relative strengths and weaknesses of each.

7. Acknowledgments

Partial funding for this project was provided by Fish, Farm, and Forest Commission (FFFC), San Francisco, CA. The author wishes to thank Ms. Nadine Wilson for her assistance composing the manuscript.

References

Agarwal, C. L. and Tikkiwal, B. D.: 1975, 'Two-stage sampling on successive occasions'. in *Proceedings of the 62nd Session of Indian Science Congress Association, Part III*. p. 31.

Bailar, B. A.: 1975, 'The effects of rotation group bias on estimates from panel surveys', *J. Amer. Statis. Assoc.* **70**, 23–29.

Bailar, B. A.: 1979, 'Rotation sample biases and their effects on estimates of change', *Bull. Intern. Statis. Inst.* **48**(2), 385–407.

Binder, D. A. and Hidiroglou, M. A. : 1988, 'Sampling in Time', in P. R. Krishnaiah and C. R. Rao (eds): *Handbook of Statistics (Vol. 6)*, Amsterdam, pp. 187–211, North Holland.

Breidt, F. J. and Fuller, W. A.: 1999, 'Design of supplemented panel surveys with application to the national resources inventory', *J. Agric. Biolog. Environ. Statis.* **4**(4), 391–403.

Cochran, W. G.: 1977, *Sampling Techniques*. New York: Wiley, 3 edition.

Duncan, G. J. and Kalton, G.: 1987, 'Issues of design and analysis of surveys across time', *Inter. Statis. Rev.* **55**, 97–117.

Eager, C., Miller-Weeks, M., Gillespie, A. J. R. and Burkman, W.: 1991, 'Summary Report: Forest Health Monitoring – New England/Mid-Atlantic', Technical Report NE-INF-115-92, U.S. Department of Agriculture, Forest Service, Radnor, PA.

Eckler, A. R.: 1955, 'Rotation Sampling', *Ann. Mathem. Statis.* **26**, 664–685.

Edwards, D.: 1998, 'Issues and themes for natural resources trend and change detection', *Ecolog. Applic.* **8**(2), 323–325.

Ernst, T. L., Leibowitz, N. C., Roose, D., Stehman, S. and Urquhart, N. S.: 1995, 'Evaluation of USEPA Environmental monitoring and assessment program's (EMAP) – Wetlands sampling design and classification'. *Environ. Manage.* **19**, 99–113.

Fuller, W. A.: 1999, 'Environmental surveys over time', *J. Agric. Biolog. Environ. Statis.* **4**(4), 331–345.

Gurney, M. and Daly, J. F.: 1965, 'A Multivariate Approach to Estimation in Periodic Sample Surveys', in *Proceedings of the American Statistical Association, Section on Survey Research Methods*, pp. 247–257.

Hansen, M. H., Hurwitz, W. N. and Madow, W. G. : 1953, *Sample Survey Methods and Theory*. New York: Wiley.

Hansen, M. H., Hurwitz, W. N., Nisselson, H. and Steinberg, J.: 1955, 'The redesign of the census current population survey', *J. Amer. Statis. Assoc.* **50**, 701–719.

Huang, E. and Ernst, L.: 1981, 'Comparison of An Alternate Estimator to the Current Composite Estimator in CPS', in *Proceedings of the Survey Research Methods Section*. pp. 303–308, American Statistical Association.

Jessen, R. J.: 1942, 'Statistical investigation of a sample survey for obtaining farm facts', *Iowa Agric. Exper. Station Res. Bull.* **304**, 54–59.

Kish, L.: 1965, *Survey Sampling*. New York: Wiley.

Kish, L.: 1983, 'Data collection for Details over Space and Time', in T. Wright (ed.): *Statistical Methods and the Improvement of Data Quality*. Orlando, FL, pp. 73–84, Academic.

Kish, L.: 1986, 'Timing of surveys for public policy', *Austr. J. Statis.* **28**, 1–12.

Krishnaiah, P. R. and Rao, C. R.: 1988, *Handbook of Statistics 6 – Sampling*. New York: North Holland.

Kumar, S. and Lee, J.: 1983, 'Evaluation of composite estimation for the Canadian Labour Force Survey', *Surv. Methodol.* **9**, 178–201.

Leatherberry, E. C., Spencer, J. J. S., Schmidt, T. L. and Carroll, M. R.: 1995, 'An Analysis of Minnesota's Fifth Forest Resources Inventory, 1990'. Technical Report Resource Bulletin NC-165, U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, MN.

McRoberts, R. E. and Hansen, M.: 1999, 'Annual forest inventories for the North Central Region of the United States', *J. Agric. Biolog. Environ. Statis.* **4**(4), 361–371.

Moisen, G. G. and Edwards, J. T. C.: 1999, 'Use of generalized linear models and digital data in a forest inventory of Northern Utah', *J. Agric. Biolog. Environ. Statis.* **4**(4), 372–390.

Nusser, S. M., Breidt, F. J. and Fuller, W. A.: 1998, 'Design and estimation for investigating the dynamics of natural resources', *Ecolog. Applic.* **8**(2), 234–245.

Nusser, S. M. and Goebel, J. J.: 1997, 'The national resources inventory: a multi-resource monitoring program', *Ecolog. Environ. Statis.* **4**, 181–204.

Olsen, A. R., Sedransk, J., Gotway, C., Liggett, W., Rathbun, S., Reckhow, K., Young, L. and Edwards, D.: 1999, 'Statistical issues for monitoring ecological and natural resources in the United States', *Environ. Monit. Assess.* **54**(1), 1–45.

Olsen, A. R. and Smith, E. P.: 1999, 'Introduction to the special issue on surveys over time', *J. Agric. Biolog. Environ. Statis.* **4**(4), 328–330.

Overton, W. S., White, D. and Stevens, D. L.: 1990, 'Design Report for EMAP, Environmental Monitoring and Assessment Program', Technical Report EPA/600/3-91/053, U.S. Environmental Protection Agency, Washington, DC.

Patterson, H. D.: 1950, 'Sampling on successive occasions with partial replacement of units', *J. Royal Statis. Soc., Series B* **12**, 241–255.

Prabhu-Ajgaonkar, S. G.: 1967, 'The theory of univariate sampling on successive occasions under the general correlation patterns', *Austr. J. Statis.* **18**, 56–63.

Rao, J. N. K. and Graham, J. E.: 1964, 'Rotation designs for sampling on repeated occasions', *J. Amer. Statis. Assoc.* **63**, 99–112.

Reams, G. A. and Deusen, P. V.: 1999, 'The southern annual forest inventory system', *J. Agric. Biolog. Environ. Statis.* **4**(4), 346–360.

Sarndal, C. E., Swensson, B. and Wretman, J. H.: 1992, *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Singh, D.: 1968, 'Estimates in successive sampling using multi-stage design', *J. Amer. Statis. Assoc.* **69**, 99–109.

Skalski, J. R.: 1990, 'A design for long-term status and trends', *J. Environ. Manage.* **30**, 139–144.

Smith, T. M. F.: 1978, 'Principles and Problems in the Analysis of Repeated Surveys', in M. K. Namboodini (ed.): *Survey Sampling and Measurement*. New York, pp. 201–206, Academic Press.

Tikkiwal, B. D.: 1979, 'Successive sampling – A review', in: *Proceedings of the 42nd Session of the International Statistical Institute held in Manilla, Book 2*. pp. 367–384.

Urquhart, N. S. and Kincaid, T. M.: 1999, 'Designs for detecting trend from repeated surveys of ecological resources', *J. Agric. Biolog. Environ. Statis.* **4**(4), 404–414.

Urquhart, N. S., Paulsen, S. G. and Larsen, D. P.: 1998, 'Monitoring for policy-relevant regional trends over time', *Ecolog. Appl.* **8**(2), 246–257.

Wolter, K.: 1979, 'Composite estimation in finite populations', *J. Amer. Statist. Assoc.* **74**, 604–613.

Yates, F.: 1949, *Sampling methods for Censuses and Surveys*, London: Charles Griffin, 1 edition.